This paper is published as part of a PCCP Themed Issue on:

Explicit-r_{12} Correlation Methods and Local Correlation Methods

Guest Editors: Hans-Joachim Werner and Peter Gill

Editorial

Papers

Density matrix renormalisation group Lagrangians

Garnet Kin-Lic Chan*

Received 28th March 2008, Accepted 25th April 2008
First published as an Advance Article on the web 20th May 2008
DOI: 10.1039/b805292c

We introduce a Lagrangian formulation of the density matrix renormalisation group (DMRG). We present Lagrangians which, when minimised, yield the optimal DMRG wavefunction in a variational sense, both within the general matrix product ansatz and within the canonical form of the matrix product that is constructed within the DMRG sweep algorithm. Some of the results obtained are similar to elementary expressions in Hartree–Fock theory, and we draw attention to such analogies. The Lagrangians introduced here will be useful in developing theories of analytic response and derivatives in the DMRG.

I. Introduction

The density matrix renormalisation group (DMRG) of White[1–3] is a recent addition to the methods of quantum chemistry.[4–10] Unlike many other correlation methods the DMRG is not based on excitations from a Hartree–Fock reference but rather on a new kind of highly flexible reference function. In quantum chemistry, it has led to advances in the treatment of strongly interacting (*i.e.* multi-reference) problems. For molecules that are large in one spatial dimension, the cost of the DMRG is only quadratic in the number of localised orbitals and it is therefore a quadratic-scaling multi-reference method for such systems.[11] We have applied a quadratic-scaling DMRG algorithm to study conjugated polymers,[12,13] light-harvesting pigments,[14] and the metal–insulator transition in hydrogen chains[11] with full treatment of multi-reference correlations in as large as 100-electron, 100-orbital complete active spaces.

Early formulations of the DMRG primarily used the language of the numerical renormalisation group that reflects the history of its development. Such language is very different from the usual language of quantum chemistry. However, as is now understood, the DMRG algorithm simply minimises the energy of a wavefunction ansatz known as the matrix product state.[3,10,15–17] This ansatz has a very different structure from most quantum chemical wavefunctions and the unique strengths and weaknesses of the DMRG method can be understood from this point of view.[10]

In a prior publication[10] we have presented an introduction to the DMRG from the wavefunction perspective. In the current work, we continue along this line of presentation and describe simple reformulations of the DMRG that connect the method with well-known *Lagrangian* techniques in quantum chemistry.[18–20] All our results are of a very elementary nature, but we feel there is sometimes value to writing out such things explicitly. In the past, Lagrangian formulations have provided a unified and systematic language by which to derive many results in the area of perturbation theory and analytic derivatives and response.[18–20] We believe that the Lagrangian formulation of the DMRG presented here will be useful in a similar way.

We start in section II by recalling the matrix product structure of the DMRG wavefunction. In section III we write down a simple Lagrangian for the DMRG wavefunction and the corresponding stationary equations. These resemble the Fock orbital equations of Hartree–Fock theory and we discuss this similarity. In DMRG calculations that are based on the traditional sweep algorithm, one implicitly uses not the most general form of the DMRG ansatz, which contains some redundancy, but rather a special *canonical* form.[10] The canonical form of the DMRG wavefunction is reviewed in the first part of section IV while in the second part we introduce the additional constraints that have to be applied to Lagrangians to ensure that the minimising wavefunction is of canonical form. We finish by demonstrating explicitly that minimising the canonical DMRG Lagrangian yields exactly the same solution conditions on the DMRG wavefunction as the original sweep algorithm of the DMRG.

II. The DMRG wavefunction

Recall the full configuration interaction expansion of the wavefunction in Fock space. In terms of Slater determinants written in the occupation number representation $|n_1 \cdots n_k\rangle$, where n_i is the occupation of orbital i taking values $0,1_\alpha,1_\beta,1_\alpha 1_\beta$, this is

$$|\Psi\rangle = \sum_{n_1 n_2 n_3 \ldots n_k} \Psi^{n_1 n_2 n_3 \ldots n_k} |n_1 n_2 n_3 \cdots n_k\rangle, \qquad (1)$$

$$\sum_i n_i = N. \qquad (2)$$

In the DMRG ansatz, the expansion coefficient $\Psi^{n_1 n_2 n_3 \ldots n_k}$ is represented by a contracted product of the tensors, where each tensor is associated with the Fock space of a given orbital. In the context of the DMRG it is more usual to refer to orbitals as sites, and thus we refer to the tensors as site-functions. Thus we write

$$\Psi^{n_1 n_2 n_3 \ldots n_k} = \psi^{n_1}_{i_1} \psi^{n_2 i_1}_{i_2} \psi^{n_3 i_2}_{i_3} \cdots \psi^{n_k}_{i_{k-1}} \qquad (3)$$

Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York, 14853, USA. E-mail: gc238@cornell.edu

where we have used the Einstein summation convention that we will employ throughout this work (*i.e.* repeated upper and lower indices are contracted). The number of coefficients in each site-function (save for the first and last) is $4M^2$, where M is the dimension of each i index. (It is conventional to take the dimension of each i index to be the same.)

From eqn (3) we see that the DMRG ansatz has a contracted matrix product structure. For this reason it is known as a matrix product state.[3,10,15,16] The matrix product maybe used to reconstruct the Slater determinant expansion of the DMRG wavefunction (see *e.g.* ref. 21). The product nature is reminiscent of the orbital product ansatz in Hartree–Fock theory. However, there are some important differences. Firstly, the number of site-functions is the size of the basis k, rather than the number of electrons N. Thus the product structure of the DMRG is expressed in the full Fock space, not in the N-particle Hilbert space. Secondly, the ansatz is a contracted product rather than a simple product. Thus correlations are introduced between the orbital Fock spaces, by virtue of the contraction structure of the i indices in the ansatz.

III. The DMRG Lagrangian

We can determine the best site-functions in the DMRG ansatz in a variational sense by minimising the energy subject to normalisation of the wavefunction. The corresponding Lagrangian is

$$\mathscr{L}[\Psi] = \langle\Psi|\hat{H}|\Psi\rangle - E\langle\Psi|\hat{1}|\Psi\rangle \tag{4}$$

To evaluate the Lagrangian explicitly with the DMRG ansatz we use a Fock representation of the Hamiltonian. The matrix elements of the Hamiltonian are written as

$$\langle n_1 n_2 \ldots n_k|\hat{H}|n'_1 n'_2 \ldots n'_k\rangle = H^{n'_1 \ldots n'_k}_{n_1 \ldots n_k} \tag{5}$$

In terms of the site-functions, the energy term $\langle\Psi|\hat{H}|\Psi\rangle$ becomes

$$\left(\psi^{n_1}_{i_1}\psi^{n_2 i_1}_{i_2}\ldots\psi^{n_k}_{i_{k-1}}\right)H^{n'_1\ldots n'_k}_{n_1\ldots n_k}\left(\psi^{i'_1}_{n'_1}\psi^{i'_2}_{n'_2 i'_1}\ldots\psi^{i'_{k-1}}_{n'_k}\right) \tag{6}$$

while the normalisation term $\langle\Psi|\hat{1}|\Psi\rangle$ is

$$\left(\psi^{n_1}_{i_1}\psi^{n_2 i_1}_{i_2}\ldots\psi^{n_k}_{i_{k-1}}\right)\left(\psi^{i'_1}_{n_1}\psi^{i'_2}_{n_2 i'_1}\ldots\psi^{i'_{k-1}}_{n_k}\right) \tag{7}$$

Note that each term in the Lagrangian is quadratic in each of the site-functions. At the minimum, the derivative of the Lagrangian with respect to the site functions vanishes. Then, the stationary equation satisfied by each site function is

$$F^{n'_p i'_{p-1} i_p}_{i'_p n_p i_{p-1}}\psi^{n_p i_{p-1}}_{i_p} = ES^{n'_p i'_{p-1} i_p}_{i'_p n_p i_{p-1}}\psi^{n_p i_{p-1}}_{i_p} \tag{8}$$

with the matrix elements of the operator $F[p]$ defined as

$$\begin{aligned} F^{n'_p i'_{p-1} i_p}_{i'_p n_p i_{p-1}} &= \left(\psi^{n_1}_{i_1}\ldots\cancel{\psi^{n_p i_{p-1}}_{i_p}}\ldots\psi^{n_k}_{i_{k-1}}\right) \\ &\quad\times H^{n'_1\ldots n'_k}_{n_1\ldots n_k}\left(\psi^{i'_1}_{n'_1}\ldots\cancel{\psi^{i'_p}_{n_p i'_{p-1}}}\ldots\psi^{i'_{k-1}}_{n'_k}\right) \end{aligned} \tag{9}$$

$$= \langle\Psi|\hat{H}|\Psi\rangle_p \tag{10}$$

and the overlap operator $S[p]$ defined as

$$\begin{aligned} S^{n'_p i'_{p-1} i_p}_{i'_p n_p i_{p-1}} &= \left(\psi^{n_1}_{i_1}\ldots\cancel{\psi^{n_p i_{p-1}}_{i_p}}\ldots\psi^{n_k}_{i_{k-1}}\right)\left(\psi^{i'_1}_{n_1}\ldots\cancel{\psi^{i'_p}_{n_p i'_{p-1}}}\ldots\psi^{i'_{k-1}}_{n_k}\right) \\ &= \langle\Psi|\hat{1}|\Psi\rangle_p \end{aligned} \tag{11}$$

Here the struck-out symbols indicate that the corresponding site-functions are omitted from the sum, and this is denoted also by the more compact expectation value notation. Considering $F^{n'_p i'_{p-1} i_p}_{i'_p n_p i_{p-1}}$, $S^{n'_p i'_{p-1} i_p}_{i'_p n_p i_{p-1}}$ and $\psi^{n_p i_{p-1}}_{i_p}$ as matrices $\boldsymbol{F}$, $\boldsymbol{S}$ and the vector $\boldsymbol{\psi}$ respectively, the stationary equations for the site-functions may be written in the form

$$\boldsymbol{F\psi} = E\boldsymbol{S\psi} \tag{12}$$

Thus, the stationary equations are analogous to the orbital Fock equations of Hartree–Fock theory[22] where each site-function is an eigenfunction of an effective site "Fock" operator $F[p]$. However, unlike in Hartree–Fock theory, the Fock operator is different for each site, and all site-functions possess the same eigenvalue E. The site Fock operator $F[p]$ may be decomposed into local-site and off-site terms. Assuming the usual form of the electronic Hamiltonian

$$\hat{H} = t^{ij}a^\dagger_i a_j + v^{ijkl}a^\dagger_i a^\dagger_j a_k a_l \tag{13}$$

where for simplicity we are assuming summations over the spin-labels of the orbitals, *i.e.* $t^{ij}a^\dagger_i a_j = t^{i\sigma j\sigma'}a^\dagger_{i\sigma}a_{j\sigma}$. We define the local-site contribution to $F[p]$ as

$$F[p]^{(\text{local})} = \langle\Psi|t^{pp}a^\dagger_p a_p + v^{pppp}a^\dagger_p a^\dagger_p a_p a_p|\Psi\rangle_p \tag{14}$$

and the off-site contributions as

$$\begin{aligned} F[p]^{(\text{off-site})} &= \langle\Psi|t^{ij}a^\dagger_i a_{j(\text{not } i=j=p)} \\ &\quad + v^{ijkl}a^\dagger_i a^\dagger_j a_k a_{l(\text{not } i=j=k=l=p)}|\Psi\rangle_p \end{aligned} \tag{15}$$

This division is analogous to the division of the Fock operator into one-electron and two-electron Coulomb-exchange terms. In particular, the off-site contributions represent the contributions of the average "field" of all the sites to the local Fock operator at site p.

IV. The canonical DMRG Lagrangian

A. Canonical form of the DMRG wavefunction

The DMRG wavefunction as written in eqn (3) possesses many redundant degrees of freedom. For example, given an arbitrary invertible matrix T, we can obtain multiple equivalent matrix product approximations for the wavefunction tensor $\Psi^{n_1\ldots n_k}$ by inserting T, T^{-1} in between two site-functions, *e.g.*

$$\begin{aligned} \Psi^{n_1\ldots n_k} &= \psi^{n_1}_{i_1}\ldots\psi^{n_p i_{p-1}}_{i_p}\psi^{n_{p+1} i_p}_{i_{p+1}}\ldots\psi^{n_k}_{i_{k-1}} \\ &= \psi^{n_1}_{i_1}\ldots\psi^{n_p i_{p-1}}_{i_p}\left(T^{i_p}_{i'_p}T^{-1}_{i''_p}{}^{i'_p}\right)\psi^{n_{p+1} i''_p}_{i_{p+1}}\ldots\psi^{n_k}_{i_{k-1}} \end{aligned} \tag{16}$$

Thus minimisation of the DMRG Lagrangian (4) does not define the site-functions uniquely, but only up to pairs of transformations.[10,14]

The original sweep algorithm used to optimise the DMRG wavefunction does, however, define a particular choice of site-functions at convergence. These site-functions are canonical in ways which resemble the properties of canonical orbitals in Hartree–Fock theory. In Hartree–Fock theory, the canonical orbitals diagonalise the Lagrange multipliers associated with orbital orthonormality. As we show in the next section, the canonical site functions obtained from the DMRG sweep algorithm diagonalise a matrix of Lagrange multipliers associated with orthogonality constraints.

Let us first recall how the sweep algorithm leads to a canonical form of the DMRG wavefunction and site-functions. We will then extract the solution conditions satisfied by the canonical site-functions at the convergence of the sweep algorithm. We will assume here some familiarity with the DMRG sweep algorithm and we refer readers to our earlier work and review for a complementary discussion.[6,10] (In particular, here we will focus on the "one-site" variant of the DMRG algorithm.[6,23,24])

In the sweep algorithm the site-functions are seen as transformation matrices which define sets of renormalised bases. For example, at block configuration $\boxed{\bullet 1 \cdots \bullet p-1}\,\bullet p\,\boxed{\bullet p+1 \cdots \bullet k}$, the first $p-1$ site-functions define many-body "left" basis functions recursively through

$$\begin{aligned} |l_{p-1}\rangle &= \sum_{n_{p-1} l_{p-2}} L_{l_{p-1}}^{n_{p-1} l_{p-2}} |n_{p-1} l_{p-2}\rangle \\ &= \sum_{n_1 \ldots n_{p-1}} L_{l_1}^{n_1} L_{l_2}^{n_2 l_1} \ldots L_{l_{p-1}}^{n_{p-1} l_{p-2}} |n_1 \ldots n_{p-1}\rangle \end{aligned} \tag{17}$$

while site-functions $p+1 \ldots k$ define many-body "right" basis functions recursively through

$$\begin{aligned} |r_p\rangle &= \sum_{n_{p+1} r_{p+1}} R_{r_p}^{n_{p+1} r_{p+1}} |n_{p+1} r_{p+1}\rangle \\ &= \sum_{n_{p+1} \ldots n_k} R_{r_p}^{n_{p+1} r_{p+1}} \ldots R_{r_{k-2}}^{n_{k-1} r_{k-1}} R_{r_{k-1}}^{n_k} |n_{p+1} \ldots n_k\rangle \end{aligned} \tag{18}$$

The transformation matrices are orthogonal in the sense that $\langle l_{p-1} | l'_{p-1}\rangle = \delta_{l'_{p-1}}^{l_{p-1}}$ and similarly for the right basis functions; this implies

$$\begin{aligned} L_{n_q l_{q-1}}^{l_q} L_{l'_q}^{n_q l_{q-1}} &= \delta_{l'_q}^{l_q} \\ R_{n_q r_q}^{r_{q-1}} R_{r'_{q-1}}^{n_q r_q} &= \delta_{r'_{q-1}}^{r_{q-1}} \end{aligned} \tag{19}$$

Using these definitions of the left and right bases $\{l_{p-1}\}$, $\{r_p\}$ as well as the basis of site p, $\{n_p\}$ the total wavefunction at the block configuration $\boxed{\bullet 1 \cdots \bullet p-1}\,\bullet p\,\boxed{\bullet p+1 \cdots \bullet k}$ is expanded as

$$\begin{aligned} |\Psi\rangle &= \sum_{l_{p-1} n_p r_p} C^{l_{p-1} n_p r_p} |l_{p-1} n_p r_p\rangle \\ &= \sum_{n_1 \ldots n_k} L_{l_1}^{n_1} \ldots L_{l_{p-1}}^{n_{p-1} l_{p-2}} \\ &\quad \times C^{l_{p-1} n_p r_p} R_{r_p}^{n_{p+1} r_{p+1}} \ldots R_{r_{k-1}}^{n_k} |n_1 \ldots n_k\rangle \end{aligned} \tag{20}$$

We see that the form of the wavefunction constructed in the sweep algorithm has a matrix product structure as in eqn (3) but has additional orthogonality constraints on the site-functions (19). Also, this wavefunction provides a special meaning to the pth site-function $C^{l_{p-1} n_p r_p}$, which appears as the set of expansion coefficients associated with the renormalised product basis $\{l_{p-1} n_p r_p\}$. We refer to the DMRG wavefunction constructed in the form (20) as the site p *canonical form*.[10,14]

It is worth emphasising explicitly at this point (as suggested by a referee) some relationships between the general matrix product state considered in section II and the canonical form just introduced. In the canonical form, all site-functions, except for a particular chosen one (at site p say), are constructed to satisfy orthogonality conditions as in (19). Now, any general matrix product state (which does not satisfy these orthogonality conditions) can *always* be transformed into canonical form, by inserting suitable pairs of transformations between the site functions as in (16) to enforce the orthogonality conditions (19). However, there must be at least one site-function which cannot be constrained to be orthogonal. This corresponds to the site-function at site p in the canonical form which contains the renormalised "state" information of the wavefunction.

We now recall how the site-functions appearing in the site p canonical form of the wavefunction (20) are actually computed in the sweep algorithm. The coefficients $C^{l_{p-1} n_p r_p}$ are obtained by solving the Schrödinger equation projected into the product basis $\{l_{p-1} n_p r_p\}$

$$\langle l'_{p-1} n'_p r'_p | \hat{H} | l_{p-1} n_p r_p\rangle C^{l_{p-1} n_p r_p} = E C^{l'_{p-1} n'_p r'_p} \tag{21}$$

These coefficients determine corresponding L, R site-functions at the same site p (*i.e.* $L_{l_p}^{n_p l_{p-1}}$ and $R_{r_{p-1}}^{n_p r_p}$) as eigenvectors of appropriate reduced density matrices. For example, $L_{l_p}^{n_p l_{p-1}}$ is obtained from the eigenvectors of a density matrix $\Gamma_{n'_p l'_{p-1}}^{n_p l_{p-1}}$ constructed by tracing over the right indices of the wavefunction coefficients

$$\Gamma_{n'_p l'_{p-1}}^{n_p l_{p-1}} = C^{l_{p-1} n_p r_p} C_{l'_{p-1} n'_p r_p} \tag{22}$$

$$\Gamma_{n'_p l'_{p-1}}^{n_p l_{p-1}} L_{l_p}^{n'_p l'_{p-1}} = w_{l_p} L_{l_p}^{n_p l_{p-1}} \tag{23}$$

while $R_{r_p}^{n_p r_{p+1}}$ is obtained from the eigenvectors of a density matrix $\Gamma_{n'_p r'_p}^{n_p r_p}$ obtained by tracing over the left indices

$$\Gamma_{n'_p r'_p}^{n_p r_p} = C^{l_{p-1} n_p r_p} C_{l_{p-1} n'_p r'_p} \tag{24}$$

$$\Gamma_{n'_p r'_p}^{n_p r_p} R_{r_{p-1}}^{n'_p r'_p} = w_{r_{p-1}} R_{r_{p-1}}^{n_p r_p} \tag{25}$$

The L and R site-functions at site p do not themselves appear in the site p canonical form; rather we need the L site functions at sites $1 \cdots p-1$ and the R site functions at sites $p+1 \ldots k$. But these can be obtained by solving the effective Schrödinger equation (21) at other block configurations in the sweep. Sweeping through block configurations $\boxed{\bullet 1 \cdots \bullet p-1}\,\bullet p\,\boxed{\bullet p+1 \cdots \bullet k}$ for $p = 1 \ldots k$, and solving for the wavefunction coefficients C at each block configuration, we can obtain all the L and R site functions appearing in the site p canonical form (20).[6,10]

Note that any wavefunction written in the canonical form of one site (say p) can always be written exactly in the canonical form of another site (say q). In this sense, canonical forms at different sites are simply different representations of the same wavefunction.[10] More precisely, given $C^{l_{q-1}n_qr_q}$, $L^{n_ql_{q-1}}_{l_q}$ at site q, we can always find $C^{l_{p-1}n_pr_p}$, $R^{n_pl_{p-1}}_{r_p}$ at site $p > q$ such that

$$\begin{aligned} &L^{n_1}_{l_1} \ldots L^{n_ql_{q-1}}_{l_q} \ldots C^{l_{p-1}n_pr_p} \ldots R^{n_k}_{r_{k-1}} \\ &= L^{n_1}_{l_1} \ldots C^{l_{q-1}n_qr_q} \ldots R^{n_pr_{p-1}}_{r_p} \ldots R^{n_k}_{r_{k-1}} \end{aligned} \tag{26}$$

In the sweep algorithm, the conversion between the canonical forms of the DMRG wavefunction at neighbouring sites is known as the wavefunction transformation,[6,10,25] and it is commonly used to accelerate the convergence of the sweeps. At convergence, if $C^{l_{p-1}n_pr_p}$ solves the effective Schrödinger equation (21) at site p, then the corresponding $C^{l_{q-1}n_qr_q}$ determined through the wavefunction transformation solves the effective Schrödinger equation (21) at site q.

Let us now summarise the solution conditions satisfied by the site-functions appearing in the site p canonical form (20) at the convergence of the DMRG sweep algorithm.

1. For a specified site (p, say), the wavefunction coefficients $C^{l_{p-1}n_pr_{p+1}}$ satisfy the effective Schrödinger equation (21) and satisfy the normalisation condition $C^{l_{p-1}n_pr_{p+1}}C_{l_{p-1}n_pr_{p+1}} = 1$,

2. The L and R site-functions are each orthogonal in the sense of (19) and are related to the C site-functions (in the corresponding canonical forms) as eigenvectors of the corresponding density matrices (23) and (25),

3. The C site-functions appearing in all the canonical forms from site $1 \ldots k$ are related through the wavefunction transformation (26).

B. Lagrangian formulation

Let us now show how the above conditions 1–3 satisfied by the canonical site-functions at the convergence of the sweep algorithm can be obtained by minimising an appropriate canonical Lagrangian. We first note that C is constrained to have unit norm while the L, R site-functions are orthogonal in the sense of (19). Thus we write a Lagrangian with these constraints

$$\begin{aligned} \mathscr{L}[\Psi] =& \langle\Psi|\hat{H}|\Psi\rangle - E(C^{l_{p-1}n_pr_p}C_{l_{p-1}n_pr_p} - 1) \\ &- \sum_{q<p} \mu^{l'_q}_{l_q}(L^{l_q}_{n_ql_{q-1}}L^{n_ql_{q-1}}_{l'_q} - \delta^{l_q}_{l'_q}) \\ &- \sum_{q>p} \mu^{r'_{q-1}}_{r_{q-1}}(R^{r_{q-1}}_{n_qr_q}R^{n_qr_q}_{r'_{q-1}} - \delta^{r_{q-1}}_{r'_{q-1}}) \end{aligned} \tag{27}$$

At the minimum, derivatives of the Lagrangian with respect to all L, C, R site-functions must vanish. Differentiating with respect to the coefficients $C^{l_{p-1}n_pr_p}$, we obtain an effective Fock eigenvalue equation for C similar to eqn (10). However, the overlap matrix in the Fock equation is simply an identity matrix, since at the solution point all L, R site-functions are orthogonal in the sense of eqn (19). By direct evaluation

$$\begin{aligned} S^{n'_pl'_{p-1}r'_p}_{n_pl_{p-1}r_p} =& (L^{n_1}_{l_1} \ldots L^{n_{p-1}l_{p-2}}_{l_{p-1}} C^{l_{p-1}n_pr_p} R^{n_{p+1}r_{p+1}}_{r_p} \ldots R^{n_k}_{r_{k-1}}) \\ &\times (L^{l'_1}_{n_1} \ldots L^{n_{p-1}l'_{p-2}}_{l'_{p-1}} C_{l'_{p-1}n_pr'_p} R^{r'_p}_{n_{p+1}r'_{p+1}} \ldots R^{r'_{k-1}}_{n'_k}) \end{aligned} \tag{28}$$

and using (19)

$$\begin{aligned} &(L^{n_1}_{l_1} \ldots L^{n_{p-1}l_{p-2}}_{l_{p-1}})(L^{l'_1}_{n_1} \ldots L^{l'_{p-1}}_{n_{p-1}l'_{p-2}}) = \delta^{l'_{p-1}}_{l_{p-1}} \\ &(R^{n_pr_{p+1}}_{r_p} \ldots R^{n_k}_{r_{k-1}})(R^{r'_{p1}}_{n'_pr'_{p+1}} \ldots R^{r'_{k-1}}_{n'_k}) = \delta^{r'_p}_{r_p} \\ &\Rightarrow S^{n'_pl'_{p-1}r'_p}_{n_pl_{p-1}r_p} = \delta^{n'_pl'_{p-1}r'_p}_{n_pl_{p-1}r_p} \end{aligned} \tag{29}$$

Thus, the site Fock equation for C is written simply as

$$F^{n_pl_{p-1}r_p}_{n'_pl'_{p-1}r'_p} C^{l'_{p-1}n'_pr'_p} = EC^{l_{p-1}n_pr_p} \tag{30}$$

Comparing this with the effective Schrödinger equation (21) we see that $F^{n_pl_{p-1}r_p}_{n'_pl'_{p-1}r'_p} = \langle l'_{p-1}n'_pr'_p|\hat{H}|l_{p-1}n_pr_p\rangle$, and thus (30) is simply the same as solution condition 1 from the sweep algorithm.

Next we consider minimising $\mathscr{L}$ with respect to the left and right site-functions. In each case there are two non-vanishing contributions to the derivative, one from the energy expression $\langle\Psi|\hat{H}|\Psi\rangle$ and the other from the orthogonality constraint. We will work out only the derivatives with respect to the left site-functions explicitly as similar expressions hold for derivatives with respect to the right site-functions. The derivative of the energy expression is

$$\begin{aligned} \partial/\partial L^{n_ql_{q-1}}_{l_q}\langle\Psi|\hat{H}|\Psi\rangle =& (L^{n_1}_{l_1} \ldots L^{n_ql_{q-1}}_{l_q} \ldots C^{l_{p-1}n_pr_{p+1}} \ldots R^{n_k}_{r_{k-1}}) \\ &\times H^{n'_1 \ldots n'_k}_{n_1 \ldots n_k}(L^{l'_1}_{n'_1} \ldots L^{l'_q}_{n'_ql'_{q-1}} \ldots C_{l_{p-1}n_pr_{p+1}} \ldots R^{r_{k-1}}_{n'_k}) \\ =& W^{n'_ql'_{q-1}l_q}_{n_ql_{q-1}l'_q} L^{l'_q}_{n'_ql'_{q-1}} \end{aligned} \tag{31}$$

while the derivative of the orthogonality constraint is

$$\begin{aligned} &-\partial/\partial L^{n_ql_{q-1}}_{l_q} \sum_{m<p} \mu^{l'_m}_{l_m}(L^{l_m}_{l_ml_{m-1}}L^{n_ml_{m-1}}_{l'_m} - \delta^{l_m}_{l'_m}) \\ &= -\mu^{l_q}_{l'_q}L^{l'_q}_{n_ql_{q-1}} \end{aligned} \tag{32}$$

and thus at the minimum, where $\partial\mathscr{L}/\partial L^{n_ql_{q-1}}_{l_q} = 0$,

$$W^{n'_ql'_{q-1}l_q}_{n_ql_{q-1}l'_q}L^{l'_q}_{n'_ql'_{q-1}} = \mu^{l_q}_{l'_q}L^{l'_q}_{n_ql_{q-1}} \tag{33}$$

Now the minimising condition (33) does not immediately resemble solution conditions 2 and 3 for the canonical site-functions from the convergence of the sweep algorithm. To demonstrate the equivalence, we first recall that any minimum of the canonical Lagrangian (27) *is also a minimum of the simple Lagrangian* (4) in section III that did not have the additional orthogonality constraints. This is because as mentioned above, we can always insert transformations as in (16)

to convert a general matrix product state (3) to a DMRG canonical form (20), and such transformations do not change the energy or wavefunction normalisation appearing in the simple Lagrangian (4). Thus, given some set of L, C, R that minimise the canonical Lagrangian (27), these all satisfy site Fock equations as in (10), but with a unit overlap matrix (from the orthogonality conditions on L, R). Then, we can substitute the Fock equation (10) in the energy derivative (31), and we find for $W^{n_q l_{q-1} l_q}_{n'_q l'_{q-1} l'_q}$

$$W^{n_q l_{q-1} l_q}_{n'_q l'_{q-1} l'_q} = E(L^{n_1}_{l_1} \dots L^{n_q l_{q-1}}_{l_q} \dots C^{l_{p-1} n_p r_p} \dots R^{n_k}_{r_k}) \times \delta^{n'_1 \dots n'_k}_{n_1 \dots n_k} (L^{l'_1}_{n'_1} \dots L^{l'_q}_{n'_q l'_{q-1}} \dots C_{l'_{p-1} n'_p r'_p} \dots R^{r'_{k-1}}_{n'_k}) \quad (34)$$

Next, we transform the wavefunctions appearing in (34) through the wavefunction transformation (26), so that the C site-function is associated with site $q+1$, *i.e.*

$$L^{n_1}_{l_1} \dots L^{n_q l_{q-1}}_{l_q} \dots C^{l_{p-1} n_p r_p} \dots R^{n_k}_{r_k} = L^{n_1}_{l_1} \dots L^{n_q l_{q-1}}_{l_q} C^{l_q n_{q+1} r_{q+1}} \dots R^{n_k}_{r_k} \quad (35)$$

Also we observe that the C site function $C^{l_q n_{q+1} r_{q+1}}$ can always be decomposed into the product of a matrix with an orthogonal matrix which we recognise as $R^{n_{q+1} r_{q+2}}_{r_{q+1}}$

$$C^{l_q n_{q+1} r_{q+1}} = c^{l_q r_{q+1}} R^{n_{q+1} r_{q+2}}_{r_{q+1}} \quad (36)$$

Finally at the minimum of the Lagrangian, all the L and R site-functions are orthogonal in the sense of (19) and thus we can evaluate the contracted products of the L site-functions and the R site-functions appearing in (34) explicitly (substituting (36) for $C^{l_q n_{q+1} r_{q+1}}$)

$$\left(L^{n_1}_{l_1} \dots L^{n_{q-1} l_{q-2}}_{l_{q-1}}\right) \delta^{n'_1 \dots n'_q}_{n_1 \dots n_q} \left(L^{l'_1}_{n'_1} \dots L^{l'_{q-1}}_{n_{q-1} l'_{q-2}}\right) = \delta^{l'_{q-1} n'_q}_{l_{q-1} n_q}$$

$$\left(R^{n_{q+1} r_{q+2}}_{r_{q+1}} \dots R^{n_k}_{r_{k-1}}\right) \delta^{n'_{q+1} \dots n'_k}_{n_{q+1} \dots n_k} \left(R^{r'_{q+1}}_{n'_{q+1} r'_{q+2}} \dots R^{r'_{k-1}}_{n'_k}\right) = \delta^{r'_{q+1}}_{r_{q+1}} \quad (37)$$

We can now use all these simplifications (35), (36) and (37) to simplify the expression for the energy derivative $W^{n_q l_{q-1} l_q}_{n'_q l'_{q-1} l'_q}$ (34). We find

$$W^{n_q l_{q-1} l_q}_{n'_q l'_{q-1} l'_q} = \delta^{n_q l_{q-1}}_{n'_q l'_{q-1}} D^{l_q}_{l'_q} \quad (38)$$

where $D^{l_q}_{l'_q}$ is a density matrix built from the coefficients $c^{l_q r_{q+1}}$

$$D^{l_q}_{l'_q} = c^{l_q r_{q+1}} c_{l'_q r_{q+1}} \quad (39)$$

and the Lagrangian minimising condition from the energy derivative (33) becomes

$$D^{l_q}_{l'_q} L^{l'_q}_{n_q l_{q-1}} = \mu^{l_q}_{l'_q} L^{l'_q}_{n_q l_{q-1}} \quad (40)$$

This minimising condition (40) is in fact a density matrix solution condition very similar to conditions 2, 3 arising from convergence of the DMRG sweep algorithm. To see the connection explicitly we recognise that the density matrix $D^{l_q}_{l'_q}$ (constructed from $c^{l_q r_{q+1}}$) is related to the density matrix in the sweep algorithm $\Gamma^{n_q l_{q-1}}_{n'_q l'_{q-1}}$ (constructed from $C^{l_{q-1} n_q r_q}$) in a simple way using eqn (36)

$$\Gamma^{n_q l_{q-1}}_{n'_q l'_{q-1}} = L^{n_q l_{q-1}}_{l_q} D^{l_q}_{l'_q} L^{l'_q}_{n'_q l'_{q-1}} \quad (41)$$

Next, we multiply (40) on both sides with the L site function

$$\left(L^{n_q l_{q-1}}_{l_q} D^{l_q}_{l'_q} L^{l'_q}_{n'_q l'_{q-1}}\right) L^{n'_q l'_{q-1}}_{l''_q} = L^{n_q l_{q-1}}_{l_q} \mu^{l_q}_{l'_q} \left(L^{l'_q}_{n'_q l'_{q-1}} L^{n'_q l'_{q-1}}_{l''_q}\right) \quad (42)$$

Substituting in (41) for the first bracketed term and using the orthogonality of the L site-functions (19) for the second bracketed term this becomes

$$E\Gamma^{n_q l_{q-1}}_{n'_q l'_{q-1}} L^{n'_q l'_{q-1}}_{l''_q} = L^{n_q l_{q-1}}_{l_q} \mu^{l_q}_{l''_q} \quad (43)$$

This is now *identical* to the density matrix eigenvector condition (33) (up to a multiplicative factor of E) if we simply perform a unitary transformation to diagonalise μ such that

$$\mu^{l_q}_{l''_q} = E w^{l_q} \delta^{l_q}_{l''_q} \quad (44)$$

Thus we have arrived at our final result: the density matrix eigenvector condition of the sweep algorithm, which defines the L and R site functions in the canonical form of the DMRG wavefunction, is equivalent to minimising the canonical Lagrangian (27), up to a unitary transformation of each site function which does not affect the energy but which diagonalises the matrix of Lagrange multipliers μ. The DMRG site-functions are thus "canonical" site functions in a way analogous to the canonical Hartree–Fock orbitals, which are similarly obtained from any energy minimising set of orbitals, by performing a unitary transformation to diagonalise the orthonormality constraints.[22]

We finish by observing that the canonical form of the DMRG wavefunction is associated with a more complicated Lagrangian than the general matrix product state. We might therefore wonder why we use the canonical form in DMRG calculations, aside from its historical link to the traditional sweep algorithm. A central characteristic of the canonical form is the conditioning of the equations when we solve for the site functions one-by-one. In the general matrix product state, the highly non-linear nature of the ansatz can be expected to lead to ill-conditioned overlap matrices in the site Fock equations, and indeed we have observed this in (unpublished) calculations using the general form of matrix product state. Considerations of orthogonality as maintained by the transformation to a canonical form at each step of the DMRG sweep are thus important to obtain numerically stable optimisation algorithms for the DMRG wavefunction.

V. Conclusions

We have shown that we can write down Lagrangians which on minimisation yield optimal density matrix renormalisation group (DMRG) wavefunctions in a variational sense. In particular, we have demonstrated the equivalence between minimising a canonical form of Lagrangian, and converging the DMRG energy through the original sweep algorithm, up to certain unitary transformations of the variational

parameters in the DMRG wavefunction which leave the Lagrangian invariant. With an increasing understanding of the DMRG from a wavefunction ansatz perspective, it is natural to look towards developing analytic derivative and response techniques as are available for other kinds of wavefunction ansatz in quantum chemistry. The results presented here are a first step in that direction.

Acknowledgements

We would like to thank the referees for their detailed comments on the manuscript. The author acknowledges support from the US Department of Energy under award #DE-FG02-07ER46432.

References

1 S. R. White, *Phys. Rev. Lett.*, 1992, **69**, 2863.
2 S. R. White, *Phys. Rev. B*, 1993, **48**, 10345.
3 U. Schollwöck, *Rev. Mod. Phys.*, 2005, **77**, 259.
4 S. R. White and R. L. Martin, *J. Chem. Phys.*, 1999, **110**, 4127.
5 A. O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri and P. Palmieri, *J. Chem. Phys.*, 2001, **115**, 6815.
6 G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.*, 2002, **116**, 4462.
7 Ö. Legeza, J. Röder and B. A. Hess, *Phys. Rev. B*, 2003, **67**, 125114.
8 G. Moritz, B. A. Hess and M. Reiher, *J. Chem. Phys.*, 2005, **122**, 024107.
9 D. Zgid and M. Nooijen, *J. Chem. Phys.*, 2008, **128**, 014107.
10 G. K.-L. Chan, J. Dorando, D. Ghosh, J. Hachmann, E. Neuscamman, H. Wang and T. Yanai, *Condens. Matter Phys.*, 2007, arXiv:cond-mat 0711.1398v1, also available in the series *Progress of Theoretical Chemical and Physics*, Springer, 2008, vol. 18, http://www.springer.com/chemistry/book/978-1-4020-8706-6.
11 J. Hachmann, W. Cardoen and G. K.-L. Chan, *J. Chem. Phys.*, 2006, **125**, 144101.
12 J. Hachmann, J. J. Dorando, M. Avilés and G. K.-L. Chan, *J. Chem. Phys.*, 2007, **127**, 134309.
13 J. J. Dorando, J. Hachmann and G. K.-L. Chan, *J. Chem. Phys.*, 2007, **127**, 084109.
14 D. Ghosh, J. Hachmann, T. Yanai and G. K.-L. Chan, *J. Chem. Phys.*, 2008, **128**, 144117.
15 S. Östlund and S. Rommer, *Phys. Rev. Lett.*, 1995, **75**, 3537.
16 S. Rommer and S. Östlund, *Phys. Rev. B*, 1997, **55**, 2164.
17 C. M. Dawson, J. Eisert and T. J. Osbourne, *Phys. Rev. Lett.*, 2008, **100**, 130501.
18 T. Helgaker and P. Jørgensen, *Theor. Chim. Acta*, 1989, **75**, 111.
19 T. Helgaker, P. Jørgensen and N. C. Handy, *Theor. Chim. Acta*, 1989, **76**, 227.
20 H. Koch and P. Jørgensen, *J. Chem. Phys.*, 1990, **93**, 3333.
21 G. Moritz and M. Reiher, *J. Chem. Phys.*, 2007, **126**, 244109.
22 A. Szabo and N. Ostlund, *Modern Quantum Chemistry*, McGraw-Hill, New York, 1989.
23 S. R. White, *Phys. Rev. B*, 2005, **72**, 180403.
24 D. Zgid and M. Nooijen, *J. Chem. Phys.*, 2008, **128**, 144115.
25 S. R. White, *Phys. Rev. Lett.*, 1996, **77**, 3633.